

AN: PAT 2001-530796  
TI: Operating method for speech-controlled device for motor vehicle involves ad hoc generation and allocation of new speech patterns using adaptive transcription  
PN: DE19942869-A1  
PD: 15.03.2001  
AB: NOVELTY - The method involves allocating speech or speech commands to an equipment function in the vehicle via speech pattern comparison. Ad hoc new functions can be facilitated by ad hoc generation and allocation of new speech patterns and these ad hoc generations are performed by adaptive transcription. DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for an arrangement for operating a speech-controlled device for a motor vehicle.; USE - For operating a speech-controlled device for a motor vehicle. ADVANTAGE - Developed to enable non-catalogued speech command functional allocations to be learned and generated automatically.  
PA: (VOLS ) VOLKSWAGEN AG;  
IN: CRULL T; WENGELNIK H;  
FA: DE19942869-A1 15.03.2001;  
CO: DE;  
IC: G10L-015/06;  
MC: W04-V01;  
DC: P86; W04;  
PR: DE1042869 08.09.1999;  
FP: 15.03.2001  
UP: 12.10.2001

---





⑮ **BUNDESREPUBLIK  
DEUTSCHLAND**



**DEUTSCHES  
PATENT- UND  
MARKENAMT**

⑫ **Offenlegungsschrift**  
⑩ **DE 199 42 869 A 1**

⑤① Int. Cl.<sup>7</sup>:  
**G 10 L 15/06**

②① Aktenzeichen: 199 42 869.7  
②② Anmeldetag: 8. 9. 1999  
②③ Offenlegungstag: 15. 3. 2001

**DE 199 42 869 A 1**

⑦① Anmelder:  
Volkswagen AG, 38440 Wolfsburg, DE

⑦② Erfinder:  
Crull, Torsten, 38106 Braunschweig, DE; Wengelnik,  
Heino, Dr., 38440 Wolfsburg, DE

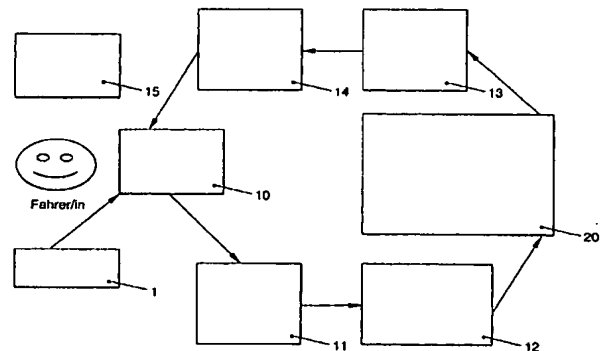
⑤⑥ Für die Beurteilung der Patentfähigkeit in Betracht  
zu ziehende Druckschriften:

DE	41 10 300 C2
DE	41 20 308 A1
DE	39 31 638 A1
EP	08 91 589 B1
EP	05 62 138 A1
WO	96 37 881 A1

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

⑤④ Verfahren und Einrichtung zum Betrieb einer sprachgesteuerten Einrichtung bei Kraftfahrzeugen

⑤⑦ Die Erfindung betrifft ein Verfahren sowie eine Einrichtung zum Betrieb einer sprachgesteuerten Einrichtung bei Kraftfahrzeugen, bei welchem/welcher die Sprache bzw. die Sprachbefehle über Sprachmustervergleich einer Gerätefunktion im Kraftfahrzeug zugewiesen wird. Um hierbei zu erreichen, daß nicht-katalogisierte Sprachbefehle und Funktionszuweisungen selbsttätig erlernbar und katalogisierbar sind, ist erfindungsgemäß vorgeschlagen, dass zusätzlich zu den vorgegebenen Sprachbefehlsauslösbaren Funktionen, durch Ad-hoc-Generierungen und -Zuweisungen von neuen Sprachmustern, ad hoc neue Funktionen ermöglicht werden, und dass diese Ad-hoc-Generierungen durch adaptive Transkription vorgenommen werden.



**DE 199 42 869 A 1**

## Beschreibung

Die Erfindung betrifft ein Verfahren sowie eine Einrichtung zum Betrieb einer sprachgesteuerten Einrichtung bei Kraftfahrzeugen, bei welchem bzw. bei welcher die Sprache bzw. die Sprachbefehle über Sprachmustervergleich einer Gerätefunktion im Kraftfahrzeug zugewiesen wird bzw. werden, gemäß Oberbegriff der Patentansprüche 1 und 11.

Einrichtungen dieser Art zur sprachgesteuerten Bedienung von Geräten innerhalb des Kraftfahrzeuges sind vielfach bekannt. Sie dienen übergeordnet dem Zweck den Fahrer nicht mehr der Notwendigkeit einer Tastenbedienung während des Fahrbetriebes zu unterwerfen und ihn damit abzulenken, sondern der Fahrzeugführer kann bei der sprachgesteuerten Befehlseingabe seinen Blick völlig unbegleitet auf den Verkehrsfluß gerichtet halten. Neben diesem enormen Sicherheitsaspekt spielt natürlich der Komfort dabei eine nicht unerhebliche Rolle.

Aus dem Stand der Technik, wie beispielsweise aus der US-Patentschrift 4 870 686, ist eine generelle sprachgenerierte Steuerung von Geräten beispielsweise eines Telefons bekannt. Hierbei werden die Sprachbefehle zunächst als Sprachnachricht erfaßt und sodann beispielsweise durch Mustervergleich ausgewertet und entsprechenden Funktionen zugewiesen.

Aus der EP 0 621 582 ebenso wie aus der EP 0 780 828 ist eine Spracherkennung durch Mustervergleich bekannt, wobei unterschiedliche Sprachmuster auch zu einem Befehl zugehörig sein können, so daß das System adaptiv eine Mustererkennung an den jeweils speziellen Bediener anpaßt.

Aus der EP 0 797 185 sind verschiedene Verfahren zur Spracherkennung bekannt. Zum einen können ausschließlich fabrikatorisch festgelegte Einzelworte zu einem so fixierten Wortschatz zusammengestellt sein, der bei der Spracherkennung verwendet wird. Dabei kann das Verfahren auf einer Phonem-Erkennung basieren.

Mit Phonem-Erkennung ist die Erkennung einzelner Laute gemeint. Hier können Verarbeitungsverfahren wie beispielsweise die Hidden-Markov-Modellierung verwendet werden. Dabei werden die von einem Benutzer eingesprochenen, digitalisierten Sprachsignale zunächst analysiert und diesbezügliche Merkmalsvektoren abgeleitet, die für die Spracherkennung wichtige Informationen des Sprachsignals beinhalten. Die dabei ermittelten Merkmalsvektoren werden ausschließlich mit für Phonem-Segmente typischen Prototyp-Merkmalsvektoren verglichen, welche beispielsweise einem dazu vorgesehenen Speicher zugeordnet werden. Weitere Verfahren können auf der Verwendung sog. Templates basieren. Hierbei basiert die Sprachbefehlerkennung auf einem Ganzwortvergleich zwischen einem eingesprochenen digitalisierten Sprachsignal und den während einer Trainingsphase eingesprochenen und zur sprecherabhängigen Spracherkennung abgespeicherten Sprechproben, den sog. Templates. Beim zweitgenannten Verfahren wird bereits klar, dass hier grundsätzlich stärker noch als bei dem oben beschriebenen Verfahren eine Festlegung von Sprachbefehlen vorher erfolgen muß. Das heißt, die Sprachbefehle können nicht ergänzt werden, es sei denn, es werden auf recht aufwendige Weise neue Templates eingelesen und bewertet, die dann auch den entsprechend neuen Funktionen zugewiesen werden müssen.

Aus der DE 197 05 471 A1 ist eine preiswerte Spracherkennung bekannt, die mittels einer Enveloppen-Bewertung der Signale ebenfalls ganze Signalpakete bewertet, wobei bei der elektronischen Verarbeitung Transversalfilter kombiniert mit Perzeptronnetzen eingesetzt werden.

Aus der EP 0 508 225 ist definitionsgemäß die Verwendung einer phonem-gestützten Spracherkennung angege-

ben. Hierbei werden zu den natürlichen Phonems noch artifizielle Phonems hinzugenommen, um eine präzisere, spezifischere Erkennung zu erhalten.

Aus der DE 195 33 541 C1 ist eine Sprachbefehlseingabe für Kraftfahrzeuge bekannt. Dabei werden adaptive Muster abgelegt, um die Sprachgewohnheiten unterschiedlicher Sprecher erkennen zu können und eine entsprechend präzisere Funktionszuweisung durchführen zu können.

Den Verfahren insgesamt ist lediglich eine vorher festgelegte Muster- und Funktionsliste zugeordnet.

Dabei werden spracheingabegesteuerte Eingaben, insbesondere beim sprachgesteuerten Betrieb der Audioanlage im Kraftfahrzeug verwendet. Hierbei ergibt sich jedoch das Problem, daß bei Fahrten auf längeren Strecken die Sender wechseln. Mit dem Wechsel der Sender wechseln natürlich auch die über RDS übermittelten Sendernamen. Herkömmliche Verfahren und Einrichtungen sind dabei nicht in der Lage, hier eine entsprechende Berücksichtigung dieses Sachverhaltes miteinzubeziehen.

Der Erfindung liegt somit die Aufgabe zugrunde, ein Verfahren sowie eine Einrichtung der gattungsgemäßen Art dahingehend weiterzubilden, daß auch nichtkatalogisierte Sprachbefehl-Funktionszuweisungen selbsttätig erlernbar und generierbar sind.

Die gestellte Aufgabe wird hinsichtlich eines Verfahrens der gattungsgemäßen Art erfindungsgemäß durch die kennzeichnenden Merkmale des Patentanspruches 1 gelöst.

Weitere vorteilhafte Ausgestaltungen der verfahrensgemäßen Erfindung wird in den Ansprüchen 2-10 angegeben. Hinsichtlich einer Einrichtung der gattungsgemäßen Art wird die gestellte Aufgabe erfindungsgemäß durch die kennzeichnenden Merkmale des Patentanspruches 11 gelöst.

Eine weitere vorteilhafte Ausgestaltung ist in Anspruch 12 angegeben.

Kern der Erfindung ist hierbei, dass zusätzlich zu vorgegebenen sprachbefehlsauslösbaren Funktionen, wie sie sozusagen katalogisiert abgespeichert sind, durch Ad-hoc-Generierung und -Zuweisung von neuen Sprachmustern, ad hoc neuen möglichen Funktionen ermöglicht werden, und dass diese Ad-hoc-Generierung durch adaptive Transkription vorgenommen wird. Dies heißt mit anderen Worten, dass neue Funktionen, wie beispielsweise neue, über das RDS der Kfz-Audioanlage erfasste Sendernamen, automatisch transkribiert werden, in Sprachmuster zu den zugehörigen Namen, und dass die zugehörigen neuen Sprachmuster nun der Liste der Sprachbefehle und der entsprechenden Funktionszuweisung hinzugefügt werden.

Somit ist es dann möglich, dass ohne weitere Programmierung, d. h. völlig selbsttätig, bei einkommendem Signal über die RDS-Codierung eines Sendernamens, der Sendernamen als solches transkribiert wird in ein Sprachmuster. Das heißt, dass der Fahrzeugführer oder der Sprachbediener nunmehr den neuen, möglicherweise vorher noch nicht in der Auflistung der bekannten Sprachbefehle enthaltenen Sendernamen sprachgesteuert eingeben kann, und die vorherige bereits erkannte Funktionszuweisung zu diesem neuen Sprachmuster lässt dann wiederum den Aufruf dieses neuen Senders im Radio des Kraftfahrzeuges sprachgesteuert zu.

In weiterer vorteilhafter Ausgestaltung werden die neuen Sprachmuster dann dauerhaft zur elektronischen Liste der Sprachbefehle und der zugehörigen Sprachmuster hinzugefügt.

Die Spracherkennung kann dabei phonem-basiert sein, womit die Auswertung einzelner Laute und nicht Gesamtworte als solches der Mustererkennverfahren zugeordnet sind. Dabei werden natürlich auch die neuen Funktionen bzw. die zu den neuen Funktionen transkribierten neuen Sprachmuster phonem-basiert generiert bzw. auch transkri-

biert. Durch die Verwendung phonem-basierter Spracherkennung ist es im übrigen auch erst möglich eine solche nachträgliche Transkription vornehmen zu können. Eine andere Transkription als phonem-basierend wäre ohnehin schwierig.

In weiterer vorteilhafter Ausgestaltung ist vorgesehen, daß eine Sprachbefehlscingabe jeweils auf ein jeweiliges Gerät einwirkt. Dies wäre beispielsweise bei einer Minimallösung der Fall, wenn z. B. nur das Autoradio sprachbefehlsgesteuert werden soll.

In weiterer vorteilhafter Ausgestaltung kann aber jedoch auch vorgesehen sein, dass eine gemeinsame Sprachbefehlscingabe interobjektiv auf verschiedene Geräte einwirkt. Das heißt, ausgehend von einer zentralisierten Sprachbefehlscingabe und Spracherkennung werden nunmehr verschiedene Geräte gesteuert. Hierzu muß dann noch außer der Auswertung des Sprachmusters und der Erkennung der Zugehörigkeit eines entsprechenden Befehles, auch noch das Gerät erkannt werden, dem diese Funktion zugeordnet ist.

In weiterer vorteilhafter Ausgestaltung ist angegeben, dass die Telefonanlage bzw. auch die Telefonanlage im speziellen, nummern- oder pseudonymwahlgesteuert wird. Damit ist gemeint, daß der Fahrzeugführer entweder die Ziffern nacheinander eingibt oder aber daß er eine auf ein Pseudonym gespeicherte Telefonnummer aus einem Telefonspeicher abrufen. Beispielsweise unter dem Wort "Büro" dann die Nummer an seinem Arbeitsplatz gewählt wird.

In weiterer vorteilhafter Ausgestaltung ist angegeben, daß die Audioanlage mit Senderkennung sprachgesteuert wird, indem die jeweils ad hoc neuen Sendernamen und dgl. transkribiert und so ansteuerbar sind.

Weiterhin kann vorgesehen sein, daß die Audioanlage mit Audiodatenträgern und den dort zugewiesenen Daten bzw. ad hoc zugewiesenen Daten transkribierbar und somit ansteuerbar sind. Damit sind die Daten gemeint, die sich auf dem Datenträger selbst befinden. Hierzu können entweder Titelziffer oder auch die Auswahl einer speziellen CD aus einem CD-Wechsler gemeint sein.

In weiterer vorteilhafter Ausgestaltung ist angegeben, daß die über Telematikdienste oder World wide web (www)-Dienste ad hoc einkommenden Daten/funktionen ad hoc in sprachbefehlsansteuerbare Sprachmuster transkribiert werden. Hiermit ist die Einbeziehung auch der genannten Telematik oder www-Dienste gemeint. Über diese Funkverbindung können Daten abgerufen werden, die für sich wiederum funktionsrelevant für verschiedene Geräte innerhalb des Fahrzeuges sind. Dazu können die Telefonanlage, aber auch die Navigationsanlage relevant sein. Auch hierbei ist es außerordentlich vorteilhaft, wenn neue einkommende Befehle sodann transkribierbar in Sprachmuster umsetzbar sind. Die so transkribierten neuen Befehle entstehen automatisch, d. h. ohne weitergehende Programmierung; d. h. den besagten neuen Funktionen sind ohne weitere Programmierung wiederum sprachgesteuert aufrufbar sind.

In weiterer vorteilhafter Ausgestaltung ist angegeben, dass Sprachmuster von extern, bspw. auch über die Telematikdienste, oder World-wide web gesendet und im Fahrzeug empfangen werden können. Auf diese Weise kann das bord-eigene System durch den besagten Zugriff nach Extern "hinzulernen".

Bezogen auf die erfindungsgemäße Einrichtung ist angegeben, daß die Sprachbefehlseinheit, eine Funktionsdeutungseinheit sowie eine Transkriptionseinheit aufweist, mit welchen ad hoc neue Funktionen erfaßt und ad hoc in Sprachmuster transkribierbar, und in der Spracheinheit als funktionszuweisbares neues Sprachbefehlsmuster abspeicherbar ist.

Die besagten einzelnen Elemente, wie Funktionsdeu-

tungseinheit und Transkriptionseinheit müssen in der entsprechend dargestellten logischen Abfolge auch chronologisch zugreifend verschaltet sein.

Vorteilhaft kann dabei sein, wenn die Sprachbefehlseinheit, mit allen oben genannten Untereinheiten wie Funktionsdeutungseinheit, Transkriptionseinheit und dgl. mehr zu einer zusammenhängenden Baugruppe zusammengefaßt sind. Über den quasi im Fahrzeug zentralisierten Betrieb einer einheitlichen Sprachbefehlseinheit sind dann eine Vielzahl von Geräten anschließbar und das System ist darüber hinaus mit der Zeit auch erweiterbar. Wichtig ist nur, daß die Kommunikation zwischen der Sprachbefehlseinheit und den einzelnen Geräten dem gleichen Protokoll unterworfen sind.

Eine Verbindung solcher Elemente untereinander innerhalb des Fahrzeuges kann dabei leicht und zweckmäßig durch die Verwendung von fahrzeuginternen Bussystemen erfolgen.

Die Erfindung ist in der Zeichnung dargestellt und nachfolgend näher beschrieben.

Das Blockschaltbild zeigt hierbei sowohl die chronologische Einbindung einzelner Elemente und ihr verfahrensgemäßes Zusammenwirken, sowie auch eine schematische Darstellung der logischen Verschaltung der erfindungsgemäßen Einrichtungselemente.

Der hier nur angedeutete Fahrer führt einen Sprachbefehl einem innerhalb des Kraftfahrzeuges angeordneten Mikrophon oder einer Mikrophananlage zu. Das Mikrophonsignal wird dabei einem Spracherkenner 10 zugeführt, der phonem-basiert ist. Phonem-basiert bedeutet, daß einzelne Laute ausgewertet werden und nicht Gesamtworte als solches. Hierdurch wird das System insgesamt selektiver und vielseitiger. Nach erfolgter phonem-basierter Spracherkennung erfolgt ein Sprachmustervergleich in einer entsprechenden Vergleichseinheit 11. Ist das entsprechende Sprachmuster erkannt und festgelegt, so erfolgt hernach innerhalb der Transformationseinheit 12 die Zuordnung des ermittelten Sprachmusters zu einer Funktion eines Gerätes.

Von der Transformationseinheit 12 aus wird dann ein entsprechendes Gerät angesteuert, was hier mit der Bezugsziffer 20 stellvertretend gekennzeichnet ist und entweder ein Telefon mit Simcard, einem Radio mit RDS, ein CD-Wechsler, eine internetfähige Kommunikationseinheit oder eine Telematikeinheit sein kann. Hierbei werden in den genannten Einheiten die entsprechenden Befehle sprachgesteuert generiert. Ebenso können die Geräte wie Telefon, Radio oder aber auch Navigationseinrichtung und dergleichen mehr, Signale von außerhalb empfangen. Hierzu können entweder abgerufene Navigationsdaten gehören oder Telematikdienst Daten oder Telefondaten oder RDS-Daten der Audioanlage.

Die empfangenen Daten können Funktionen entsprechen. Tun sie dies, so wird eine solche Funktion oder ein Sendename oder dergleichen erkannt, und hernach in einer Transkriptionseinheit 13 wird der Funktion ein Sprachmuster zugeordnet. Dieses Sprachmuster wird sodann nach erfolgter Transkription in der Transkriptionseinheit 13 der Sprachmusterspeichereinheit 14 zugeführt. Dort erfolgt die Hinzufügung als Sprachmuster zu einem funktionszugewiesenen Befehl. An der Stelle schließt sich wiederum der Kreis und der Spracherkenner 10 nimmt diesen zusätzlichen Sprachbefehl bzw. das dazugehörige Sprachmuster auf.

Insgesamt ist die Anlage somit in der Lage, neue Funktionen zu "lernen" und das dazugehörige Sprachmuster ohne vorherige Verfügbarkeit in einem Speicher selbsttätig neu zu programmieren. Dies ergibt einen erheblichen Gewinn im Bedienungskomfort, zum anderen erhöht es auch die Sicherheit, weil der Fahrzeugführer in bester Weise unterstützt wird, ohne ihn vom Fahrgeschehen abzulenken.

1. Verfahren zum Betrieb einer sprachgesteuerten Einrichtung bei Kraftfahrzeugen, bei welchem/welcher die Sprache bzw. die Sprachbefehle über Sprachmustervergleich einer Gerätefunktion im Kraftfahrzeug zugewiesen wird bzw. werden, **dadurch gekennzeichnet**, daß zusätzlich zu den vorgegebenen sprachbefehlsauslösbaren Funktionen, durch Ad-hoc-Generierungen und -Zuweisungen von neuen Sprachmustern, ad hoc neue Funktionen ermöglicht werden, und dass diese Ad-hoc-Generierungen durch adaptive Transkription vorgenommen werden. 5
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die neuen Sprachmuster zur elektronischen Befehlsliste der Sprachbefehle und der dazugehörigen Sprachmuster hinzugefügt werden. 10
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die Spracherkennung phonem-basiert ist. 15
4. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß die aus den neuen Funktionen transkribierten Sprachmuster phonem-basiert generiert bzw. transkribiert werden. 20
5. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß eine Sprachbefehls-eingabe jeweils auf ein jeweiliges Gerät einwirkt. 25
6. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß eine gemeinsame Sprachbefehlseingabe interobjektiv auf verschiedene Geräte einwirkt. 30
7. Verfahren nach Anspruch 5 oder 6, dadurch gekennzeichnet, daß die Telefonanlage nummern- oder pseudonymwahlgesteuert wird.
8. Verfahren nach Anspruch 5 oder 6, dadurch gekennzeichnet, daß die Audioanlage mit Senderkennung sprachgesteuert wird, indem die jeweils ad hoc neuen Sendernamen oder dergleichen transkribiert und so ansteuerbar sind. 35
9. Verfahren nach Anspruch 5 oder 6, dadurch gekennzeichnet, daß die Audioanlage mit Audiodatenträgern und den dort zugewiesenen Daten bzw. die ad hoc zugewiesenen Daten transkribierbar und so ansteuerbar sind. 40
10. Verfahren nach Anspruch 5 oder 6, dadurch gekennzeichnet, daß die über Telematikdienste oder world wide web (www)-Dienste ad hoc einkommenden Daten/Funktionen ad hoc in sprachbefehlsansteuerbare Sprachmuster transkribiert werden. 45
11. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß über Telematikdienste und/oder world wide web(www)-Dienste Sprachmuster zur Sprachbedienung übersendet werden. 50
12. Einrichtung zum Betrieb einer sprachgesteuerten Einrichtung bei Kraftfahrzeugen, bei welchem/welcher die Sprache bzw. die Sprachbefehle über Sprachmustervergleich in einer entsprechend elektronischen Spracheingabeeinheit, einer Gerätefunktion im Kraftfahrzeug zuweisbar ist, dadurch gekennzeichnet, daß die Sprachbefehlseinheit eine Transformationseinheit (12) sowie eine Transkriptionseinheit (13) aufweist, mit welchen ad hoc neue Funktionen erfaßt und ad hoc in Sprachmuster transkribierbar und in der Sprachmusterspeichereinheit (14) als funktionszuweisbare neue Sprachbefehlsmuster abspeicherbar sind. 60
13. Einrichtung nach Anspruch 11, dadurch gekennzeichnet, daß alle der Sprachbefehlseinheit zugeordneten Untereinheiten (10, 11, 12, 13, 14) zu einer zusammengefaßten Baugruppe zusammengefaßt sind. 65

- Leerseite -

